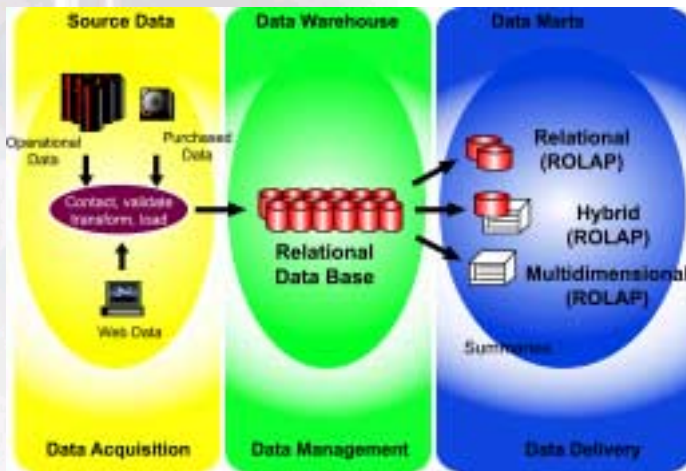# iSeries and Business Intelligence:
# Key Attributes and Considerations

***Intro***

Business Intelligence is a broad term relating to applications designed to gather, consolidate, cleanse and analyze data for purposes of understanding and acting on the key metrics that drive profitability in an enterprise. Examples include:

- Corporate Performance Measurement reporting, including Balanced Scorecard reporting and standard financial reports.
- Budget and planning applications that simplify the process and insure integrity of data during each planning cycle.
- Analytical CRM (Customer Relationship Management) applies BI practices to the marketing function to optimize budgets, execute and track campaign strategies, and generate more profit from marketing activities.
- Sales Analysis to track sales vs. forecast, build plans based on profit or packaging strategies, and nurture premier customers with better service through a more informative sales force.
- Even Human Resource departments leverage BI applications to monitor and analyze payroll classifications, training programs or labor utilization.

Armed with timely, intelligent information that is easily understood (because it is delivered in business terms), the Business Analyst is enabled to affect change and develop strategies to drive higher profits. The following picture shows a typical data warehouse foundation for Business Intelligence applications. The applications provide tools to automate processes including data acquisition, data storage and management, and data delivery.



Key attributes and considerations for choosing the IBM eServer iSeries for Business Intelligence include:

- Sophisticated Cost Based Optimization Technologies built into DB2

The sophisticated cost based **query optimizer** built into DB2 UDB for iSeries is the basis of achieving optimal performance in a BI application. The optimizer provides the "brains" behind the SQL processing which most BI applications depend on. The optimizer's goal is to build the best plan for accessing data, which will minimize extraneous I/O and therefore provide optimal performance. DB2 UDB for iSeries'

cost-based optimizer uses statistics stored automatically in the database as well as other information, such as system configuration, SMP settings, and available indexes in its development of the access plan.

IBM recently enhanced DB2 UDB for iSeries even further, introducing new optimization technology called Self-Optimizing Query Environment, including a new SQL Optimization Engine and a new statistics engine. The following chart shows the kind of improvements SQE is bringing to customers. Here's what one prominent Business Intelligence application provider stated: *"Our testing shows that IBM's DB2 UDB for iSeries is significantly faster, in some cases up to 400% faster, with the new SQL Query Engine,"* said *Mark LaRow, Vice President of Products at MicroStrategy Incorporated. "We're excited that our joint iSeries and MicroStrategy 7i customers can take full advantage of these new enhancements."*

### Complex Queries
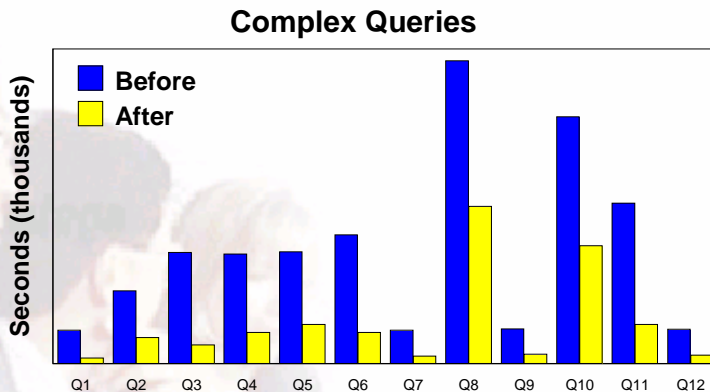


- DB2 UDB for iSeries lowers Total Cost of Ownership compared to other solutions

The eServer iSeries' unique architecture, including Single Level Storage and a tightly integrated object based operating system, have been the foundation for its leadership in Total Cost of Ownership for many years. In addition, DB2 UDB for iSeries leverages those architectural benefits by reducing the tasks required by a database administrator (DBA) typical in other architectures. For example, with DB2 UDB for iSeries:

- ➢ Data Partitioning for performance: **NOT REQUIRED**
- ➢ Moving data or indexes to avoid disk hot-spots: **NOT REQUIRED**
- ➢ Re-balancing Indexes: **NOT REQUIRED**
- ➢ Monitoring table spaces, log buffers, lock contention buffers: **NOT REQUIRED**
- ➢ Running integrity checkers or statistics collection routines: **NOT REQUIRED**

*Support, development and maintenance costs as well as our time to market would be double or triple what they are now if we had taken the Oracle route! The DB2 UDB on iSeries platform allows us to provide an extremely cost effective solution we wouldn't be able to offer on other platforms." – Don Klabunde, CTO, 360 Group, a marketing services ASP*

- Dynamic Logical Partition provides On-Demand flexibility to optimize resources

**Logical Partitioning (LPAR)** is a huge advantage for leveraging a common architecture. Logical Partitioning allows you to dynamically allocate resources across multiple copies of OS/400 within a single system footprint. LPAR allows you to carve up the system into partitions, and then dynamically allocate resources (CPU, I/O, and memory) across those partitions with a few simple clicks of a mouse.

LPAR can be applied to a Business Intelligence environment in several critical areas. First off, the ability to move resources around can significantly reduce the Extract/Transformation/Load (ETL) and calculation times typically required in a data warehouse architecture. These workloads, typically running at night, can steal processor and memory from operational application partitions that are going into batch or quiescing entirely. Once the ETL work is completed, those resources can be automatically moved back into the production partition for total optimization with minimal costs.

The following picture provides a schematic of how you could use Logical Partitioning in a BI environment:
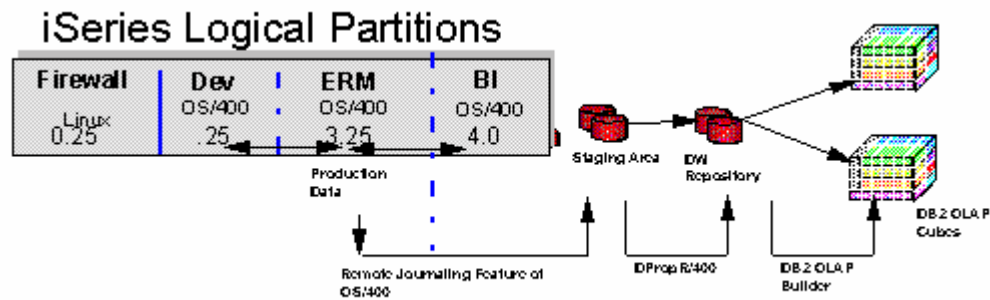


- Scalability of DB2

A key consideration in any database platform in support of BI applications is its ability to scale while maintaining acceptable performance. DB2 UDB for iSeries offers a variety of techniques to address scalability/performance. **Parallel database** operations through the **Symmetric Multiprocessing (SMP) feature of OS/400** provide the ability to split single database tasks/requests across multiple processors within a single iSeries system or logical partition (LPAR). The SMP feature provides significant performance enhancements to database queries, loads, index builds and other operations commonly used by BI applications.

A key performance technology introduced in DB2 UDB for iSeries is IBM's patented **Encoded Vector Indexing (EVIs)**. EVIs offer dramatic advancement to bitmap indexing technology. Using this advanced indexing feature, DB2 UDB for iSeries was able to run a query across a 225 GB table in a little over 35 seconds. This same query previously took over two hours.

**"The DB2 UDB enhancements, especially to the new query engine are very exciting and will enable us to continue to deliver technology enhancements to our customers, ensuring the iSeries remains a tremendous platform for building scalable, reliable, high-performance data warehouses." -** Alan Jordan, Vice President, Coglin Mill

- Leverage data transport efficiencies through Remote Journaling or High Availability offerings

With DB2 UDB for iSeries there are some fundamental approaches that can make your life easier. With V4R2, IBM introduced Remote Journaling. With Remote Journaling, you can capture changed production data and have the system route the changed data logs (journal receivers) to another iSeries logical partition (LPAR) or system.

To read the database changes from the journal receivers, you could use a low cost utility called Data Propagator to add the changed records to the reporting repository (on this 2$^{nd}$ partition/system). This essentially eliminates a key issue with any BI application - the impact of extraction or data transport on the production environment.

A side benefit of this is the ability to use the 2$^{nd}$ partition/system for other purposes as well, such as backing up the databases regularly without impacting production, or to create a **High Availability (H/A)** scenario.

- Leverage tooling or application tiers running in Linux, AIX or Windows environments

Logical Partitioning (LPAR) described above, also supports Linux and AIX operating systems. For example, multiple OS/400 and Linux partitions can be defined to a single eServer iSeries system, reducing complexity, leveraging dynamic allocation of resources and other benefits such as virtual I/O, high speed interconnects between the partitions, and disk virtualization.

How does this relate to support of Business Intelligence applications? Suppose you are using an ETL tool whose TRANSFORMATION component does not support OS/400 but can run in Linux, AIX or Windows. By partitioning or using IXA/IXS, the code logic remains within the framework of the iSeries, meaning no data has to be moved around in a network and high speed bus communications can be leveraged to optimize the entire ETL process. All data – source databases, target databases, and any intermediate data storage areas, all remain on iSeries' highly reliable controlled disk.

- Reduce Risk

Business Intelligence Applications can be a very powerful enabler for any company looking to gain insights into their data. Risk in any BI project is significantly reduced if you can leverage attributes common with current operational systems. These attributes include sharing of common hardware and software components, leveraging existing operational and administrative skills and policies, and being able to optimize investments across both operational and business intelligence applications through sharing resources and skills.

The IBM eServer iSeries is an excellent platform for minimizing risk and costs associated with implementation of this powerful technology. For additional information, please refer to:

DB2 UDB for iSeries Home Page:
http://www.ibm.com/servers/eserver/iseries/db2